THE RECEPTIVE VOCABULARY SIZE OF SPANISH 5TH GRADE PRIMARY SCHOOL STUDENTS IN CLIL AND NON-CLIL INSTRUCTION

Andrés Canga Alonso

Universidad de La Rioja

Abstract

One of the key factors at early stages of L2 learning is the number of words learners know. However, there is a paucity of research regarding the receptive vocabulary size of learners involved in Content and Language Integrated Learning (CLIL) programmes. In light of this lack of research, the 2,000 word frequency band of the Vocabulary Levels Test (VLT) was administered to explore the receptive vocabulary knowledge of 5th grade primary Spanish students learning English in two different types of instruction i.e. CLIL and non-CLIL. Sex-based differences among the participants will also be explored. Our findings show that there are statistically significant differences in favour of CLIL learners. Regarding sex and regardless of the type of instruction, the differences among boys and girls are not statistically significant. However, our results reveal that CLIL boys attain the highest scores in the VLT, being them significantly higher than the outcomes achieved by non-CLIL boys and girls. These data evince that CLIL programmes may have a positive impact for vocabulary learning.

Resumen

Uno de los factores clave en el aprendizaje temprano de una L2 es el número de palabras que saben los aprendices. Sin embargo, hay una carencia de investigación en lo referente al tamaño del vocabulario receptivo de los estudiantes que participan en programas de Aprendizaje Integrado de Lenguas y Contenidos (AICLE). A la luz de esta falta de investigación se administró el *Vocabulary Levels Test* (VLT) correspondiente a las 2000 palabras más frecuentes del inglés para explorar el vocabulario receptivo de estudiantes españoles de 5º de primaria en dos tipos diferentes de instrucción AICLE y no-AICLE. Asimismo, se analizarán las diferencias entre los grupos de acuerdo con el sexo de los participantes. Los resultados muestran que existen diferencias significativas a favor de los estudiantes AICLE. En lo que respecta al sexo, no se ponen de manifiesto diferencias estadísticamente significativas entre los chicos y las chicas si no tenemos en cuenta el tipo de instrucción. Sin embargo, se constata que los chicos AICLE obtienen los mejores resultados en el VLT, siendo éstos significativamente mejores que los alcanzados por los chicos y las chicas no-AICLE. Estos datos ponen de manifiesto que los programas AICLE pueden tener un impacto positivo en el aprendizaje del vocabulario.

Submission: 24/10/2014- Acceptance: 03/11/2014 ES 36 (2015): 63-85

Keywords: CLIL, Non-CLIL, VLT, receptive vocabulary, sex, word estimates.

Palabras clave: AICLE, no-AICLE, VLT, vocabulario receptivo, sexo, estimación de palabras.

1. INTRODUCTION

Vocabulary is a crucial aspect in second language acquisition (SLA), overall proficiency, and general academic achievement (Daller, van Hout and Treffers-Daller 2003; Morris and Cobb 2004). Thus, the better the lexical competence of the learners is, the better their general linguistic competence can be assumed to be. In this sense, examining learners' word knowledge can provide with interesting and reliable insights into their overall language knowledge. This can be a good way of exploring the effects of CLIL instruction in the learning of English as a foreign language (FL) by Spanish primary school students. The present paper has as its main purpose to compare the receptive vocabulary size of traditional 5th grade EFL (non-CLIL) and CLIL primary school learners in Spain, and to analyse if there are statistically significant differences regarding the receptive vocabulary knowledge of male and female learners. We are interested in examining the impact of CLIL tuition on learners' receptive lexical competence versus that of general instruction in the FL. Hence, we review the main studies dealing with receptive vocabulary size and sex-based differences. We also explore the relationship between CLIL and FL vocabulary. Report of the study conducted with its methodology, main results found and interpretation of the same follows. We conclude pointing out some lines for further research trying to overcome the main limitations of the present study.

2. CLIL AND VOCABULARY LEARNING

In recent decades, the priority given to multilingualism by European institutions favoured Content and Language Integrated (CLIL) programmes to spread in Europe. As Ruiz de Zarobe (2011) states, the acronym CLIL (Content and Language Integrated Learning) has been used as a generic term to describe all types of approaches in which a second language is used to teach certain subjects in the curriculum other than language lessons. The essence of CLIL is integration with a dual focus: "language learning is included in content classes (e.g. maths, history, geography... etc), and content from subjects is used in language learning classes"

(Mehisto, Marsh and Frigols 2008: 11). CLIL also provides real and relevant input for the learner. This input refers to the content that the teacher is presenting as well as the language for classroom management necessary to ensure that learning takes place (Muñoz 2007; Coyle, Hood and Marsh 2010).

Assuming that in CLIL settings it is necessary to progress systematically in pupils' content and language learning and use, vocabulary knowledge is of paramount importance in order to favour communication in the classroom. Feedback is also integrated into classroom discourse to encourage interaction among apprentices. Thus, the challenge in the CLIL setting is that trainees need to engage in dialogic interactions by using the vehicular language. As a result of this interaction, Dalton-Puffer (2007, 2008) reports that there are some areas where clear gains are observed in CLIL classrooms such as e.g. receptive skills, vocabulary, morphology, and creativity.

This relates to the distinction purported by Cummins' (2008) between *basic interpersonal communicative skills* (BICs) and *cognitive academic language proficiency* (CALP) as an attempt for students to catch up with their peers in academic aspects of the school language vocabulary. Thus, BICs and CALPs should be taken into account in order to draw educators' attention to the timelines and challenges that second language learners encounter. As will be illustrated in the procedure and data gathering section, the 2,000 frequency band of Vocabulary Levels Test (VLT) measures students' ability to recognise words from these lists which relates to Cummins' concept of CALP, since learners should be able to recognise and understand the meaning of the terms provided and relate them to their definitions.

As mentioned in the introduction, vocabulary knowledge is acknowledged to be of paramount importance to facilitate students' interaction in the foreign language. Researchers have tackled the issue concerning the number of words necessary to understand spoken discourse (Nation 2001; Adolphs and Schmitt 2004) and to read and comprehend texts in the native and foreign language (Anderson and Freebody 1981; Laufer 1997). Among the former researchers, Adolphs and Schmitt (2004) estimate that, at least, 2,000 word forms have to be mastered in order to understand around 90% and 94% of spoken discourse in different contexts. Among the latter, Laufer (1992, 1997) states that a text coverage of 95% can be reached with a 5,000-word English vocabulary or 3,000 word families, which agrees with the assertions made by Hazenberg and Hulstijn (1996), Nation (1993, 2001) and Cobb and Horst (2004). More recently, Nation (2006) asserts that 8,000 to 9,000 word families are needed for understanding a written text and a vocabulary of 6,000 to 7,000 word families for comprehension of spoken text, if 98% coverage of a text is desired. Hirsh and Nation (1992) also point out that knowledge of 5,000 word families is necessary to enjoy reading. As we have seen, estimates based on word frequency criteria have been calculated and research claims that gaining command of the 2,000-3,000 most frequent words as soon as possible is vital for the language learner to communicate

orally and in written form in the foreign language (Nation 1993; Nation and Waring 1997). The sooner the most frequent words are learned by students, the better their language performance will be. As Schmitt (2000: 137) claims: "The learning of these basic words cannot be left to chance, but should be taught as quickly as possible, because they open the door of further learning".

In recent decades, a considerable number of studies have investigated receptive vocabulary size or the number of words a learner knows. Most studies coincide in indicating that vocabulary size grows as proficiency level in the foreign language (Barrow et al. 1999; Fan 2000), exposure to the target language (Schmitt 1998; Golberg et al. 2008) or frequency of input (Vermeer 2001) increase. Moreover, this gain follows a systematic order related to frequency, since at the lowest levels of proficiency learners are familiar with the most frequent words, but as their experience with the foreign language increases, less frequent words are incorporated into the lexicon (Barrow et al. 1999; Vermeer 2001; Milton 2009). The probability of a word being known by foreign language learners rises with its frequency, so higher-frequency words have a greater possibility of being known. It seems evident that a content-based approach provides more opportunities to learn either explicitly or implicitly target vocabulary in meaningful situations (Muñoz 2007; Pérez-Vidal 2009) since learners are exposed to the target language for a longer period than students' enrolled in EFL classrooms. Xanthou (2010) states that CLIL has a positive impact in a group of primary school children in Cyprus regarding students` vocabulary tests results which demonstrates that by attaching words to their surroundings, the likelihood of comprehension and retention is increased. These gains in receptive vocabulary size are in line with other research conducted in Spain (Jiménez Catalán, Ruiz de Zarobe and Cenoz 2006; Jiménez Catalán and Ruiz de Zarobe 2009), where significant results were obtained in favour of the CLIL group in receptive vocabulary size. Finally, Sylven (2004, 2006) carried out two studies in Sweden who proved a positive correlation between more hours of instruction, greater vocabulary acquisition and higher degree of communicative competence in the foreign language in favour of CLIL students.

Table 1 presents a summary of previous estimates of receptive vocabulary size of L2 learners of English at primary and secondary level in CLIL and non-CLIL classrooms after having received a similar amount of hours of instruction to the students analysed in the present sample (400-800 hours). Studies are ordered according to the receptive vocabulary size of learners. As can be seen, the results obtained show considerable differences in receptive vocabulary knowledge on the part of the learners who were investigated. Thus, as could be expected, CLIL learners (Jiménez Catalán and Ruiz de Zarobe 2009) obtained better results than students of their same age in Spain with lower time of exposure to the target language (Agustín Llach and Terrazas Gallego 2012), but lower scores if compared to secondary school students in Spain and abroad (Takala 1984; Takala 1985; Milton and Meara 1998;

Terrazas Gallego and Agustín Llach 2009). This variation in the number of words learners know should be taken with caution due to differences concerning pupils, the tests which were administered and their contexts of learning. However, we will find it useful to compare the results presented in the table with the ones obtained by our students since, to our knowledge, there are no studies in Spain which have related CLIL and non-CLIL receptive vocabulary sizes at primary school level apart from Jiménez Catalán and Ruiz de Zarobe's (2009). This research was conducted in two different autonomous communities in the north of Spain (The Basque Country and La Rioja) with last year primary students whereas the present study was carried out in the same autonomous community with one year younger (5th grade primary) students from a similar background. Both studies also differ on the number of the content subjects that were taught through English since their students learnt Arts and Crafts and Sciences through English and ours received instruction only in Sciences.

Study	**Receptive Vocabulary Size**	**Hours of Instruction**	**L1**	**Participants learning context**	**Test**
Milton and Meara (1998)	1,680 words	660	Greek	Secondary School	LLEX Lingua Vocabulary Tests
Takala (1984, 1985)	1,500 words	450	Finnish	Last of Finnish Comprehensive School	Translation of Finnish words into English
Milton and Meara (1998)	1,200 words	400	German	Secondary School	LLEX Lingua Vocabulary Tests
Jiménez Catalán and Ruiz de Zarobe (2009)	800 words	960	Spanish	Primary Education (6th Grade/CLIL)	2kVLT
Terrazas Gallego and Agustín Llach (2009)	817 words	734	Spanish	Secondary Education (1st ESO/7th Grade)	2k VLT
Jiménez Catalán and	559 words	419	Spanish	Primary Education (4th	2kVLT

Terrazas Gallego (2008)				Grade)	
Agustín Llach and Terrazas Gallego (2012)	663 words	629	Spanish	Primary Education (6th Grade)	2k VLT

Table 1. Average receptive vocabulary size.

3. VOCABULARY LEARNING AND SEX

The role of sex has also occupied an outstanding place in current research on vocabulary acquisition. Receptive and productive vocabulary knowledge of male and female learners has been widely examined, and scholars have reached different conclusions. Boyle (1987) concludes that, exceptionally, boys are superior to girls in the comprehension of heard vocabulary. Similarly, Scarcella and Zimmerman (1998) find that men performed significantly better than women in a test of academic vocabulary recognition, understanding and use. In Lynn et al. (2005) and Edelenbos and Vinjé (2000) males also outperform females in vocabulary knowledge in the foreign language. By contrast, in Nyikos' study (1990) women perform better than men in a memorisation test of German vocabulary. Nevertheless, Jiménez Catalán and Terrazas Gallego (2008) discover no significant sex-based differences in performance on a receptive vocabulary test implemented with primary students. In a recent longitudinal study on receptive vocabulary knowledge and sex-based differences, Agustín Llach and Terrazas Gallego (2012) obtain similar results since they found very slight differences among males and females across grades in the context of Spanish primary education concerning their receptive vocabulary knowledge. Contrariwise, highly significant differences are found in favour of females in the mean number of words produced in response to the 15 cues of a lexical availability test (Jiménez Catalán and Ojeda Alba 2009). A set of recent studies compiled in Jimenez Catalán (2010) also point to mixed results on gender differences or tendencies. As Sunderland (2010) claims, a careful analysis of this compilation throws the conclusion that the relationships between vocabulary and sex are not enduring, but may be context and test type-specific, being also influenced by L1, age or L2 proficiency.

Considering the aforementioned studies, we can state that there is a scarcity of research with regard to receptive vocabulary size and sex-based differences in CLIL and non-CLIL instruction in Spanish primary education. For these reasons, this study aims at investigating the receptive vocabulary knowledge of 5th grade primary CLIL male and female learners to compare their receptive vocabulary size with the receptive vocabulary knowledge of a sample of non-CLIL learners of their same age and social background to state if there are significant differences in the estimation of words according to type of instruction (CLIL vs. non-CLIL) and sex. Results will be related to students' ability to understand spoken and written discourse in English (Nation 1993, Laufer 1992; Hazenberg and Hulstijn 1996; Laufer 1997; Nation 2001; Adolphs and Schmitt 2004; Cobb and Horst 2004). Hence, we set out to find answers to the following research questions:

RQ1: What is the receptive vocabulary size of 5th grade primary school learners in CLIL?

RQ2: Are there any significant differences according in students' receptive vocabulary sizes according to type of instruction (CLIL/non-CLIL?

RQ3: Can we account for significant differences between boys and girls in both types of instruction?

4. METHOD

4.1. PARTICIPANTS

The sample for the study is constituted by 255 students. The CLIL group is made up of 61 (35 boys and 26 girls) 5th grade primary students in a CLIL programme from the same school. The non-CLIL group comprises 194 learners (106 boys and 88 girls) enrolled in English as a curricular subject in the same grade of primary education. The sample was homogeneous as regards social environment since all the students lived in the same town, but non-CLIL learners were from different private and public schools in the area. Students also shared Spanish as their mother tongue (L1) and they were 10-11 years old. The groups differ in the kind of instruction they receive, i.e. CLIL vs. non-CLIL, and consequently, in the number of hours of exposure to English FL. Learners in the non-CLIL group are exposed to English through the English FL school subject, exclusively. However, learners in the CLIL group receive, apart from the weekly EFL lessons, input in English in the school subject Natural Sciences, which is taught through the medium of English not

only the amount but also the nature of the input differs between the traditional and the CLIL group. Thus, traditional learners have received approximately between 105-110 hours of exposure to EFL on a yearly basis since 1 st grade of primary. The CLIL group has received these 105-110 hours plus 72-74 more hours in CLIL science, also since 1st year of primary education.

Table 2 illustrates the approximate number of hours of exposure students have received by the time of data collection

		Hours of exposure	
Grade	Age	CLIL	Non- CLIL
5th Primary	10-11	839	524

Table 2. Hours of exposure to English FL (accumulated exposure)

4.2. PROCEDURE AND DATA GATHERING

The 2,000 word frequency-band (2k) from the receptive version of the Vocabulary Levels Test (VLT) was used to measure the receptive vocabulary size of these subjects (Schmitt, Schmitt and Clapham 2001, version 2). As mentioned in section 2, foreign language learners need to gain a command of the 2,000 most frequent words in English as soon as possible in order to be able to understand spoken and written discourse in this language (Nation 1993, Laufer 1992; Hazenberg and Hulstijn 1996; Laufer 1997; Nation 2001; Adolphs and Schmitt 2004; Cobb and Horst 2004). For this reason we decided to implement the 2k VLT, since this test is based on the frequency lists collected by West (1953) in the General Service List and the Thorndike and Lorge (1944) list, which were checked against the list compiled by Kucera and Francis (1967), known as the Brown Corpus.

In the 2k VLT (see Appendix I), test-takers have to match a target word with the corresponding definition. A total of 60 target words are used for testing. Ten groups of six words and three definitions make up the test. Each correct answer, i.e. matching each target word with its definition is given one point, so that the maximum score of the test is 30 points. The research studies that have reported on the validity and reliability of the 2k VLT (Beglar and Hunt 1999; Read 2000) evince that the test is not only valid and consistent in its measurements, but also that, in fact, it measures what it sets out to measure.

Data were collected in one session during class time. The time allotted to complete the task was 10 minutes. At the beginning of the test, clear instructions were given both orally and in written form in the students' mother tongue to clarify what they were being asked to do.

In order to calculate students´ word estimates, Nation's formula "Vocabulary size = N correct answers multiplied by total N words in dictionary (the relevant word list) divided by N items in test" (Nation 1990: 78) was applied. The data was also analysed with SPSS 19 to check whether there were statistically significant differences according to type of instruction and sex.

5. RESULTS

In this section we will present the overall results obtained by the CLIL group in the 2k VLT (RQ1). Afterwards, CLIL and non-CLIL learners will be compared in order to establish if there are statistically significant differences between both groups according to type of instruction (CLIL/non-CLIL) (RQ2). Finally, sex-based differences will be tackled in order to ascertain if they are significant for both samples of learners (RQ3).

As for the results obtained by the CLIL group in the 2k VLT (RQ1), figure 1 illustrates that the maximum score was nineteen points out of thirty which was attained by two of the students whereas the minimum score (one out of thirty) was achieved by two participants. Their mean scores (10.44) indicate that students are way back from having learnt the 2,000 most frequent words according to the frequency lists collected by Thorndike and Lorge (1944), West (1953) and Kucera and Francis (1967). Our findings also show that 11.48% scored between zero and five points, 36.06% scored between six and ten points, 40.98% of the students scored between eleven and fifteen points and 11.48% scored between sixteen and nineteen.

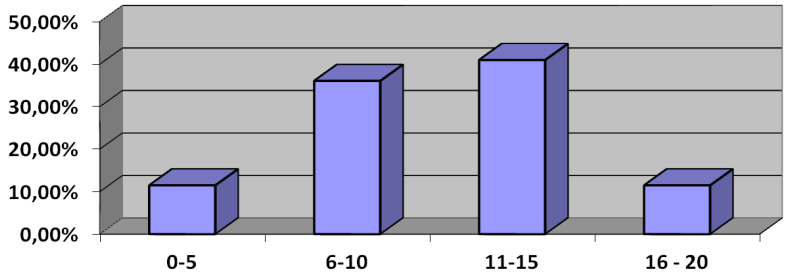

Figure 1. Frequency distribution of test scores in the CLIL group.

Students' scores were translated into a number of known words for each frequency level applying Nation's formula: "Vocabulary size = N correct answers multiplied by total N words in dictionary (the relevant word list) divided by N items in test" (1990: 78). The means obtained (696 words) confirms our previous presupposition that CLIL learners in this sample were way behind 1,000 most frequent words in English. This result is also a bit lower if compared to previous studies conducted with secondary school students in Spain and abroad since 400-700 hours of instruction led to vocabulary sizes of around 1,000 words (Takala 1984; Takala 1985; Staehr 2008; Terrazas Gallego and Agustín Llach 2009). This hypothesis is reinforced by the fact that only 11.48% of the participants have a receptive vocabulary size equal or higher than 1,000 words in the 2k VLT.

If we compare the data with the results obtained by the informants enrolled in the non-CLIL group, Table 3 shows that non-CLIL learners obtained better results when compared to CLIL students regarding maximum scores. Thus, non-CLIL maximum score was twenty (1333 words) whereas CLIL students achieved nineteen points (1267 words). As for the mean scores, CLIL learners outperformed their non-CLIL counterparts being it 10.44 (696 words) as opposed to 7.44 (499 words) for the non-CLIL group. Regarding minimum scores, there were two CLIL students who scored one point (67 words) while all the non-CLIL students, except for one informant, achieved more than six points (400 words) in the test. The outcomes reveal that CLIL learners obtained higher mean scores (696 words) if they are compared to our non-CLIL informants (499 words).

Type of instruction	Overall results				Word estimates			
	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD
CLIL (n=61)	1	19	10.4	4.5	67	1267	696	280.9
Non-CLIL (n=194)	1	20	7.5	4.1	67	1333	499	273.3

Table 3. Overall results and word estimates CLIL and non-CLIL learners

Shapiro Wilk and Kolmogorov-Smirnov tests were implemented in order to ascertain whether our sample met the normality assumption. The p-values (see table 4) indicate that we can accept normality in the CLIL group since their results are higher than 0.05, but we cannot accept it in the non-CLIL sample as they are below this figure, henceforth we have to apply non-parametric tests.

VLT		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Instruction	Statistic	gl	Sig.	Statistic	gl	Sig.
	CLIL	.089	61	.200*	.979	61	.361
	Non-CLIL	.085	194	.002	.970	194	.000

Table 4. Parametric tests: Types of Instruction

The U Mann-Whitney test was conducted to test inferential statistical differences among the CLIL and non-CLIL groups and its results reveal significant gender differences in favour of the CLIL learners at a significance level of 5% in receptive vocabulary size estimations (p=0.00). Table 5 offers these results:

VLT	
Mann-Whitney	3603
Wilcoxon W	22518
Z	-4.612
P (two-tailed)	0.00

Table 5. Results of inferential statistics according to type of instruction.

As for sex differences in CLIL instruction, CLIL boys and girls obtained the same maximum and minimum scores in the test. However, the means (10.91 vs. 9.81) is higher for boys, which indicates that their overall scores were higher. Figure 2 compares graphically the scores attained by male and female learners in the CLIL sample. As can be seen, boys attained higher scores since more than 48% scored between 11-15 points in the VLT whereas 46.16% of the girls achieved 6-10 points in the test.

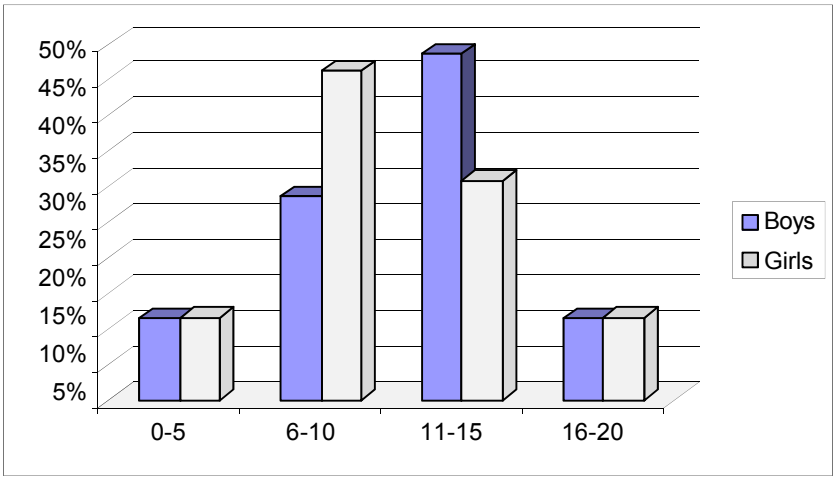

Figure 2. Frequency distribution of test scores for CLIL boys and girls.

Regarding word estimates and sex-based differences in CLIL, table 6 illustrates that mean values are higher for boys (728 words vs. 654 words for girls), whereas minimum (67 words) and maximum (1267 words) word estimates values remain the same for both sexes. As shown in figure 2 above, the higher frequency of boys' tests scores accounts for the difference in boys' and girls' mean values.

As far as sex-based differences in receptive vocabulary size in the CLIL and non-CLIL groups are concerned (see table 6), descriptive results reveal the maximum values were higher for non-CLIL female participants (1333words/20 points in the test) closely followed by CLIL boys and girls (1267 words/19 points in the test). Nevertheless, CLIL boys attained the highest means (728 words) closely followed by CLIL girls (654 words). The non-CLIL group achieves the opposite results since the mean values are slightly higher in favour of females (510 words).

Type of instruction	Sex	Overall Results				Word estimates			
		Min.	Max.	Mean	SD	Min.	Max.	Mean	SD
CLIL	Boy(n=35)	1	19	10.9	3.9	67	1267	728	260.9
	Girl (n=26)	1	19	9.8	4.6	67	1267	654	305.8
Non-CLIL	Boy (n=106)	1	16	7.3	4.0	67	1067	490	273.2
	Girl (n=88)	1	20	7.7	4.1	67	1333	510	274.4

Table 6. Overall results and word estimates according to sex and type of instruction.

Shapiro Wilk and Kolmogorov-Smirnov were used to ascertain if the sample of boys and girls, regardless of their type of instruction, met the normality assumption. As illustrated in table 7, the results obtained revealed that we could not accept normality for any of the groups, accordingly our data was analysed with the U Mann-Whitney test.

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Sex	Statistic	gl	Sig.	Statistic	gl	Sig.
VLT	Boys	.088	142	.009	.967	142	.002
	Girls	.116	113	.001	.969	113	.010

Table 7. Normality tests sex differences.

The U Man Whitney test confirmed that the p-value (p=0.67) was much higher than 5%, hence we can conclude that sex differences among the groups regarding word estimates are not statistically significant in this sample. The boxplot in figure 3 represents graphically the spread of the distribution of boys' and girls' word estimates. As can be seen, boys' word estimates are slightly higher than girls' regardless of the type of instruction (CLIL vs. non-CLIL).

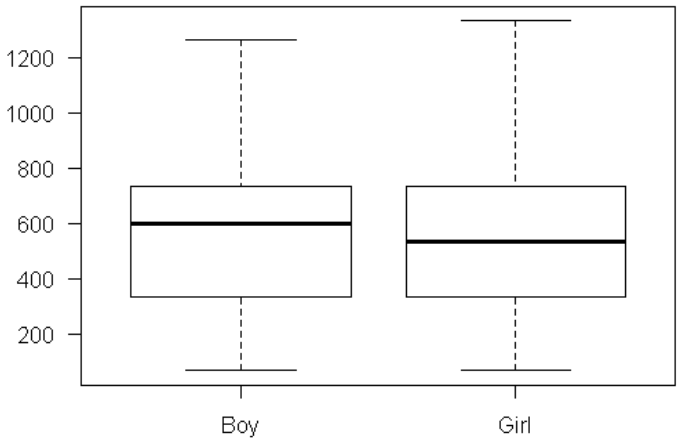

Figure 3. Sex-based differences CLIL and non-CLIL groups.

In the light of these results, we also considered relevant to compare sex and type of instruction, i.e. CLIL boys and CLIL girls, CLIL boys with non-CLIL boys and non-CLIL girls and CLIL girls with non-CLIL boys and girls to ascertain if any of the sex groups performed significantly better than the others in the 2kVLT.

Shapiro Wilk and Kolmogorov-Smirnov tests were implemented in order to ascertain whether our CLIL sample met the normality assumption. The p-values obtained (p=0.2) were higher than 5%, so the sample could be consider normal. Therefore, we applied a t-test. The p-value was much higher than 5% (p=0.31), therefore we cannot establish statistically significant differences according to sex in the CLIL group.

According to the results presented above, we believed that CLIL boys had performed significantly better in the 2k VLT than non-CLIL boys and girls. The statistical analysis implemented corroborated our presupposition, table 8 offers these results:

	CLIL boys vs, Non-CLIL boys	CLIL boys vs, Non-CLIL boys
Mann-Whitney	993	767
Wilcoxon W	6771	4595

Z	-4.17	-4.11
P (two-tailed)	0.00	0.00

Table 8. Inferential statistics CLIL boys vs. Non-CLIL boys and girls.

The results of our CLIL girls were also compared with our non-CLIL sample, and as shown in table 9, we can account for statistical significant differences among these groups of learners:

	CLIL girls vs, Non-CLIL boys	CLIL girls vs, Non-CLIL boys
Mann-Whitney	986	791
Wilcoxon W	6764	4619
Z	-2.3	-2.32
P (two-tailed)	0.021	0.02

Table 9. Inferential statistics CLIL girls vs. Non-CLIL boys and girls.

6. DISCUSSION

The findings of the present study reveal that regardless of the type of instruction (i.e. CLIL/non-CLIL) students' receptive vocabulary knowledge is lower than 1,000 words. This finding is not surprising, since it is in line with previous results with young learners and a limited amount of exposure (Webb and Chang 2012; Nurweni and Read 1999). However, these results are poorer if compared to those obtained by learners who have received similar amounts of instruction in EFL (non-CLIL) instruction in other countries (Takala 1985; Milton and Meara 1998) or even in a similar context (Terrazas Gallego and Agustín Llach 2009). However, our CLIL informants obtained better results than one year older students in the same socio-cultural context (Agustín Llach and Terrazas Gallego 2012), hence CLIL instruction seems to be beneficial.

These results might be related to the kind of vocabulary input students are exposed to in their CLIL classroom and in their textbooks. Therefore, in line with the research conducted by Jiménez Catalán and Mancebo Francisco (2008) with EFL textbooks, further research is needed to find out the type of vocabulary input which is included in the textbook teachers and students often use in the CLIL classroom as well as the number of occurrences of the words contained in them, so that students can progressively acquire new words related to the topics they are dealing with in their content lessons. In addition, the wordlists used by the 2k VLT to measure students' receptive vocabulary size (Thorndike and Lorge 1944; West 1953; Kucera and Francis 1967) are not adapted either to EFL or CLIL instruction, which might imply that the words learnt by the CLIL informants in their content lessons (Science) may not be reflected in the 2,000 most frequent English words. As an avenue for further research, future studies could test learners knowledge of specific vocabulary related to, e.g. the field of science (the CLIL subject). We could speculate that results would favour of the CLIL group, whose vocabulary size might be bigger in this semantic field. However, this is just mere speculation since we have not conducted such a test and further research is needed, which explores words students may have acquired in the content classes; and compare their results to non-CLIL learners.

Our results also prove that the students in the present study will find it difficult to understand written and spoken discourse in English since they need to master at least 2,000 word forms to be able to understand around 90% and 94% of spoken discourse in different contexts (Nation 2001; Adolphs and Schmitt 2004) and about 5,000-word English vocabulary or 3,000 word families to reach a text coverage of 95% (Laufer 1992; Hazenberg and Hulstijn 1996; Laufer 1997; Nation 2001; Adolphs and Schmitt 2004; Cobb and Horst 2004) and their means of word estimates are clearly lower than 1,000 words. Nevertheless, the CLIL learners in our sample would have an easier time watching TV programs (Webb and Rogers 2009a), or movies (Webb and Rogers 2009b), since the first 1,000 most frequent words make up for around 85% of the total word coverage and our CLIL students have gained around 70% (699 words) of the 1,000 most frequent words in the foreign language.

As regards type of instruction i.e. CLIL, non-CLIL, our findings reveal that there are statistically significant differences in favour of CLIL learners. This result is in line with previous studies, which show that CLIL or longer FL exposure programs foster vocabulary acquisition, and that benefits start cropping up after some time (Sylven 2004, 2006; Jiménez Catalán, Ruiz de Zarobe and Cenoz 2006; Dalton-Puffer 2007; Lasagabaster 2008; Jiménez Catalán and Ruiz de Zarobe 2009; Celaya and Ruiz de Zarobe 2010; Xanthou 2010). The longer exposure to English input CLIL learners have received can help explain this advantage in general receptive vocabulary size. Furthermore, the different nature of their exposure, in which traditional EFL instruction is combined with a more meaningful and contextualized content

instruction through the L2 might also account for this difference. However, this hypothesis needs to be tested with a wider sample of CLIL learners.

In short, CLIL seems to favour receptive vocabulary knowledge, but there might be other factors influencing students' receptive vocabulary acquisition, such as for example exposure time. It is indeed extremely difficult to discern whether the CLIL factor or the longer exposure (more instruction hours) is to be made responsible for CLIL advantage in our data. However, previous studies have found that increasing exposure leads to no lexical gains in young learners but that as learners grow older they benefit more from longer exposure time (Miralpeix 2007, Agustín Llach and Terrazas-Gallego 2012). In this vein, learners are also exposed to a foreign language outside the classroom context (e.g. watching videos, talking to friends through social media or even playing online games) and this exposure could also affect their vocabulary learning. It could even be considered a new way of CLIL learning since they are in contact with specific content words from the semantic fields of the topics aforementioned. Further research in this respect is warranted. We are also led to think that although CLIL learners do not learn many more words than traditional EFL learners, it can be reasonable to believe that due to the kind of exposure they have received, they might have a greater depth of knowledge than the traditional learners. However, further research is needed to investigate this.

As for sex-based differences, our data do not reveal statistically significant differences between boys and girls regardless of their type of instruction. Nevertheless, our results indicate that CLIL boys performed significantly better than non-CLIL boys and girls. They also obtained slightly higher results than our CLIL girls but differences between these two groups are non-significant. CLIL girls are also significantly better than non-CLIL students regardless of their sex. Henceforth, CLIL tuition seems to have favoured receptive vocabulary acquisition. These outcomes concur with research conducted abroad were boys attained higher results in vocabulary tests than girls (Boyle 1987; Scarcella and Zimmermann 1998), but differ from other studies carried out in a similar context and age group (Jiménez Catalán and Terrazas Gallego 2008; Agustín Llach and Terrazas Gallego (2012) where girls slightly outperformed boys. Consequently, we can conclude that the relationships between vocabulary and sex are not enduring but may related to the type of instruction i.e. CLIL/non-CLIL though these assumption should be taken with caution since the sample of CLIL learners is quite small. Thus, further research needs to be carried out with a wider sample of CLIL learners with the same amount of hours of instruction in different schools in the same area to test if the results resemble the ones obtained in the present study or contrariwise we can establish significant differences between boys' and girls' receptive vocabulary knowledge in CLIL tuition.

7. CONCLUSION

Under the light of our results, we conclude that CLIL learners have statistically significant larger receptive vocabulary sizes than non-CLIL students of their same age in a similar socio-cultural context. However, CLIL and non-CLIL receptive vocabulary knowledge is much lower than 1,000 words, which implies that they can find difficulties to understand written and spoken discourse in English. On the other hand, we find no statistically significant sex-based differences among boys and girls notwithstanding the type of instruction. On the contrary, CLIL boys attained highly significant scores in the 2kVLT than non-CLIL girls and boys and the same applies for CLIL girls when they are compared with non-CLIL learners. These results lead us to think that CLIL tuition with its contextualized, meaningful FL teaching and use and its higher number of FL exposure is a beneficial approach for vocabulary acquisition.

Our findings cannot be generalised since the study presents limitations as the number of CLIL students who took part in the study is quite reduced. It is worth mentioning that researchers find difficult to gather data from bigger samples of CLIL students since not many schools are implementing CLIL programmes in the area where this research was conducted when compared to the number of institutions whose syllabi are based on traditional EFL (non-CLIL) instruction. In this sense, further studies are necessary, which explore CLIL benefits for vocabulary acquisition for longer periods of time going into the compulsory secondary education stage, therefore longitudinal studies with CLIL learners will also be helpful to analyse is receptive vocabulary is incremental along primary and secondary school instruction. In addition, this study has only been concerned with receptive vocabulary of the most frequent words; we are inclined to believe though, that productive vocabulary and lexical knowledge of less frequent words can also derive benefits from an educational approach of the content-based type. One further limitation of the present study is the use of a single, and somewhat limited, instrument to measure vocabulary size. Using other tests for vocabulary knowledge, such as lexical availability tests might throw even more insightful results and reveal more qualitative data concerning learners' vocabulary knowledge. Further research is called for to overcome these limitations.

REFERENCES

Adolphs, Svenja, and Norbert Schmitt. "Vocabulary Coverage According to Spoken Discourse Context." *Vocabulary in a Second Language*. Eds. Paul Bogaards and Batia Laufer. Amsterdam: John Benjamins, 2004: 39-52.

Agustín Llach, Maria Pilar, and Melania Terrazas Gallego. "Vocabulary Knowledge Development and Gender Differences in a Second Language." *ELIA* 12 (2012): 45-75.

Anderson, Richard C., and Peter Freebody. "Vocabulary Knowledge." *Comprehension and Teaching: Research Reviews*. Ed. John Guthrie. Newark: International Reading Association, 1981: 203-226.

Barrow, Jack, Y. Nakanishi, and H. Ishino. "Assessing Japanese College Students' Vocabulary Knowledge with a Self-Checking Familiarity Survey." *System* 27 (1999): 223-247.

Beglar, David, and Alan Hunt. "Revising and Validating the 2000 Word Level and University Word Level Vocabulary Tests." *Language Testing* 16 (1999): 131-162.

Boyle, Joseph P. "Sex Differences in Listening Vocabulary." *Language Learning* 37.2 (1987): 273-284.

Celaya, Mª Luz, and Yolanda Ruiz de Zarobe. "First Languages and Age in CLIL and non-CLIL Contexts." *International CLIL Research Journal*, 1.3 (2010). URL: http://www.icrj.eu/13/article6.htm. 17/09/2013.

Cobb, Tom, and Marlise Horst. "Is there Room for an Academic Word List in French." *Vocabulary in a Second Language*. Eds. Paul Bogaards and Batia Laufer. Amsterdam: John Benjamins, 2004: 15-38.

Coyle, Do, Philip Hood and David Marsh. *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press, 2010.

Cummins, James. "BICS and CALP: Empirical and Theoretical Status of the Distinction." *Encyclopaedia of Language and Education, 2nd Edition, Volume 2*. Eds. Brian V. Street and Nancy H. Hornberger. New York: Springer Science + Business Media LLC, 2008: 71-83.

Daller, Helmut, Roeland Van Hout and Jeanine Treffers-Daller. "Lexical Richness in the Spontaneous Speech of Bilinguals." *Applied Linguistics* 24.2 (2003): 197-222.

Dulton-Puffer, Christiane. *Discourse in Content and Language Integrated Learning (CLIL)*. Amsterdam/Phildelphia: John Benjamins, 2007.

----. "Outcomes and Processes in Content and Language Integrated Learning (CLIL): Current Research from Europe." *Future Perspectives for English Language Teaching*. Eds. Werner Delanoy and Laurenz Volkmann. Heidelberg: Carl Winter, 2008: 139-157.

Edelenbos, Peter, and Maria P. Vinjé. "The Assessment of a Foreign Language at the End of Primary (Elementary) Education." *Language Testing* 17.2 (2000): 144-162.

Ellis, Rod. *The Study of Second Language Acquisition*. Oxford: Oxford University Press, 1995.

Fan, May. "How Big is the Gap and How to Narrow It? An Investigation into the Active and Passive Vocabulary Knowledge of L2 Learners." *RELC Journal* 31 (2000): 105-119.

Golberg, Heather, Johanne Paradis and Martha Crago. "Lexical Acquisition over Time in Minority First Language Children Learning English as a Second Language." *Applied Psycholinguistics* 29 (2008): 41–65.

Hazenberg, Susan, and Jan H. Hulstijn. "Defining a Minimal Receptive Second-Language Vocabulary for Non-Native University Students: An Empirical Investigation." *Applied Linguistics* 17 (1996): 145–63.

Hirsh, David, and Paul S. N. Nation. "What Vocabulary Size Is Needed to Read Unsimplified Texts for Pleasure?" *Reading in a Foreign Language* 8 (1992): 689-696.

Jiménez Catalán, Rosa María, ed. *Gender Perspectives on Vocabulary in Foreign and Second Language*. Basingstoke: Palgrave, 2010.

Jiménez Catalán, Rosa María and Rocío Mancebo Francisco. "Vocabulary Input in EFL Textbooks."*RESLA* 21 (2008): 147-165.

Jiménez Catalán, Rosa Maria and Julieta Ojeda Alba. "Girls' and Boys' Lexical Availability in EFL."*International Journal of Applied Linguistics* 158 (2009): 57-76.

Jiménez Catalán, Rosa María, and Yolanda Ruiz de Zarobe. "The Receptive Vocabulary of EFL Learners in Two Instructional Contexts: CLIL versus non-CLIL." *Content and Language Integrated Learning. Evidence from Research in Europe*. Eds. Rosa María Jiménez Catalán and Yolanda Ruiz de Zarobe. Bristol: Multilingual Matters, 2009: 81-92.

Jiménez Catalán, Rosa María, Yolanda Ruiz de Zarobe and Jasone Cenoz. "Vocabulary Profiles in English as a Subject and as a Vehicular Language." *VIEWZ* 15.3 (2006): 23-27. URL: http://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/views15_3_clil_special.pdf. 03/09/2013.

Jiménez, Rosa María, and Melania Terrazas Gallego. "The Receptive Vocabulary of English Foreign Language Young Learner." *Journal of English Studies* 5 (2008): 173-191.

Kucera, Henry, and Winthrop Nelson Francis. *A Computational Analysis of Present Day American English*. Providence, Rhode Island: Brown University Press, 1967.

Lasagabaster, David. "Foreign Language Competence in Content and Language Integrated Courses." *The Open Applied Linguistics Journal* I (2008): 31-42. URL:http://www.atriumlinguarum.org/contenido/Lasagabaster_TOALJ.pdf. 01/08/2013.

Laufer, Batia. "How Much Lexis Is Necessary for Reading Comprehension." *Vocabulary and Applied Linguistics*. Eds. Pierre J.L. Arnaud and Henri Béjoint. London: Macmillan, 1992: 126-132.

----. "The Lexical Plight in Second Language Reading." *Second Language Vocabulary Acquisition*: Eds. James Coady and Thomas Huckin. Cambridge: Cambridge University Press, 1997: 20-34.

Lightbown, Patsy M., and Nina Spada. "Learning English as a Second Language in a Special School in Quebec." *The Canadian Modern Language Review* 53.2 (1997): 315-355.

Lynn, Richard, David M. Fergusson and L. John Horwood. "Sex Differences on the WISC-R in New Zealand." *Personality and Individual Differences* 39 (2005): 103-114.

Mehisto, Peeter, David Marsh and Mª Jesús Frigols. *Uncovering CLIL: Content and Language Integrated Learning in Bilingual and Multilingual Contexts*. Oxford: Macmillan, 2008.

Milton, James. *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters, 2009.

Milton, James, and Paul Meara. "Are the British Really Bad at Learning Foreign Languages?" *Language Learning Journal* 18 (1998): 68-76.

Miralpeix, Inma. "Lexical Knowledge in Instructed Language Learning: The Effects of Age and Exposure." *International Journal of English Studies* 7.2 (2007): 61-84.

Morris, Lori, and Tom Cobb. "Vocabulary Profiles as Predictors of the Academic Performance of Teaching English as a Second Language Trainees." *System* 32.1 (2004): 75-87.

Muñoz, Carmen. "CLIL. Some Thoughts on its Psychological Principles." *RESLA* (2007): 11-17.

Nation, Paul S. N. *Teaching and Learning Vocabulary*. New York: Newbury, 1990.

----. "Using Dictionaries to Estimate Vocabulary Size: Essential, but Rarely Followed, Procedures." *Language Testing* 10 (1993): 27-40.

----. *Learning Vocabulary in Another Language*. New York: Cambridge University Press, 2001.

----. "How Large a Vocabulary Is Needed for Reading and Listening?" *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 63 (2006): 59-81.

Nation, Paul S. N., and Robert Waring. "Vocabulary Size, Text Coverage and Word Lists." *Vocabulary: Description, Acquisition and Pedagogy*. Eds. Norbert Schmitt and Michael McCarthy. New York: Cambridge University Press, 1997: 6-19.

Nurweni, Ari, and John Read. "The English Vocabulary Knowledge of Indonesian University Students." *English for Specific Purposes* 18.2 (1999): 161-175.

Nyikos, Martha. "Sex-related Differences in Adult Language Learning: Socialization and Memory Factors." *Modern Language Journal* 74 (1990): 273-287.

Pérez-Vidal, Carmen. The Integration of Content and Language in the Classroom: A European Approach to Education (the Second Time around)." *CLIL across Educational Levels*. Eds. Emma Dafouz and Michele C. Guerini. Madrid: Richmond, 2009: 3-16.

Read, John. *Assessing Vocabulary*. Cambridge: Cambridge University Press, 2000.

Ruiz de Zarobe, Yolanda. "Which Language Competencies Benefit from CLIL? An Insight into Applied Linguistics Research." *Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts.* Eds. Yolanda Ruiz de Zarobe, Juan Manuel Sierra and Francisco Gallardo del Puerto. Bern: Peter Lang, 2011: 123-153.

Scarcella, Robin, and Cheryl Zimmerman. "Academic Words and Gender. ESL Student Performance on a *Test* of Academic Lexicon." *Studies in Second Language Acquisition* 20 (1998): 27-49.

----. "Tracking the Incremental Acquisition of Second Language Vocabulary: A Longitudinal Study." *Language Learning* 48 (1998): 281-317.

Schmitt, Norbert. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press, 2000. Schmitt, Norbert, Diane Schmitt and Caroline Clapham. "Developing and Exploring the Behaviour of Two New Versions of the Vocabulary Level Test." *Language Testing* 18 (2001): 55-88.

Staehr, Lars Stenius. "Vocabulary Size and the Skills of Listening, Reading and Writing." *Language Learning Journal* 36 (2008): 139-52.

Sunderland, Jane. "Theorizing Gender Perspectives in Foreign and Second Language Learning." *Gender Perspectives on Vocabulary in Foreign and Second Languages*. Ed. Rosa Maria Jiménez Catalán. Basingstoke: Palgrave, 2010: 1-22.

Sylven, Liss Kristin. *Teaching English or English Teaching? On the Effects of Content and Language Integrated Learning on Swedish Learners' Incidental Vocabulary Acquisition*. Göteborg: Department of English, Göteborg University, 2004.

----. "Extramural Exposure to English among Swedish School Students Used in the CLIL Classroom?" *VIEWZ* 15.3 (2006): 47-54. URL:

http://anglistik.univie.ac.at/fileadmin/user_UniversityPressload/dep_anglist/weitere_UniversityPressloads/Views/views15_3_clil_special.pdf. 16/08/2013.

Takala, Sauli. *Evaluation of Students' Knowledge of English Vocabulary in the Finnish Comprehensive School* (Reports of the Institute of Educational Research 350). Jyväskylä: Finland, 1984.

----. "Estimating Students' Vocabulary Sizes in Foreign Language Teaching. Practice and Problems in Language Testing." *Practice and Problems in Language Testing*. Eds. Viljo Kohonen, Hikka van Essen and Christine Klein-Braley. Tampere: Finnish Association for Applied Linguistics, 1985: 157-65.

Terrazas Gallego, Melania, and María Pilar Agustín Llach. "Exploring the Increase of Receptive Vocabulary Knowledge in the Foreign Language: A Longitudinal Study." *IJES* 9.1 (2009): 113-133. URL: http://revistas.um.es/ijes/article/viewFile/90681/87481. 15/09/2013.

Thorndike, Edward, and Irvine Lorge. *The Teacher's Word Book of 30,000 Words*. New York Teachers College: Columbia University, 1944

Vermeer, Anne. "Breadth and Depth of Vocabulary in Relation to L1/L2 Acquisition and Frequency of Input." *Applied Psycholinguistics* 22 (2001): 217-34.

Webb, Stuart Alexander, and Anna C-S Chang. "Second Language Vocabulary Growth." *RELC Journal* 43.1 (2012): 113-126.

Webb, Stuart, and Michael P.H Rodgers. "The Vocabulary Demands of Television Programmes." *Language Learning* 59.2 (2009a): 335-366.

----. "The Lexical Coverage of Movies." *Applied Linguistics* 30.3 (2009b): 407-427.

Xanthou, Maria. "Current Trends in L2 Vocabulary Learning and Instruction. Is CLIL the Right Approach?" *Advances in Research on Language Acquisition and Teaching: Selected Papers* (2010): 459-471.
URL:http://www.enl.auth.gr/gala/14th/Papers/English%20papers/Xanthou.pdf. 30/09/2013.

West, Michael: *A General Service List of English Words*. London: Longman, 1953.

How to cite this article:

Canga Alonso, Andrés. "The Receptive Vocabulary Size of Spanish 5th Grade Primary School Students in CLIL and Non-CLIL Instruction." *ES. Revista de Filología Inglesa* 36 (2015): 63-85

Author's contact: andres.canga@unirioja.es

APPENDIX I. VOCABULARY LEVELS TEST 2,000 (Schmitt. Schmitt and Clapham 2001)

2,000 WORD LEVEL TEST

CENTRO_____

CURSO_____**GRUPO**_____ **FECHA**_____

APELLIDOS _____ **NOMBRE** _____

Este es un test de vocabulario. En la parte izquierda te presentamos grupos de seis palabras inglesas y a su derecha, los significados de sólo tres de ellas. **Escribe** junto a éstos, el **número** de la palabra inglesa correspondiente a dichos significados. Observa el siguiente ejemplo:

EJEMPLO		RESPUESTA CORRECTA	
1 business		1 business	
2 clock	____ part of a house	2 clock	__6__ part of a house
3 horse	____ animal with 4 legs	3 horse	__3__ animal with 4 legs
4 pencil	____ something used for writing	4 pencil	__4__ something used for writing
5 shoe		5 shoe	
6 wall		6 wall	

1 coffee		1 adopt	
2 disease	_____ money for work	2 climb	_____ go up
3 justice	_____ a piece of clothing	3 examine	_____ look at closely
4 skirt	_____ using the law in the right way	4 pour	_____ be on every side
5 stage		5 satisfy	
6 wage		6 surround	
1 choice		1 bake	
2 crop	_____ heat	2 connect	_____ join together
3 flesh	_____ meat	3 inquire	_____ walk without purpose
4 salary	_____ money paid regularly for doing a job	4 limit	_____ keep within a certain size
5 secret		5 recognize	
6 temperature		6 wander	
1 cap		1 burst	
2 education	_____ teaching and learning	2 concern	_____ break open
3 journey	_____ numbers to measure with	3 deliver	_____ make better
4 parent	_____ going to a far place	4 fold	_____ take something to someone
5 scale		5 improve	
6 trick		6 urge	
1 attack		1 original	
2 charm	_____ gold and silver	2 private	_____ first
3 lack	_____ pleasing quality	3 royal	_____ not public
4 pen	_____ not having something	4 slow	_____ all added together
5 shadow		5 sorry	
6 treasure		6 total	
1 cream		1 ancient	
2 factory	_____ part of milk	2 curious	_____ not easy
3 nail	_____ a lot of money	3 difficult	_____ very old
4 pupil	_____ person who is studying	4 entire	_____ related to God
5 sacrifice		5 holy	
6 wealth		6 social	